View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Computer Engineering and Applications Journal (ComEngApp, Universitas Sriwijaya)

Computer Engineering And Applications

Computer Engineering and Applications Vol. 8, No. 2, June 2019

PCA-Based on Feature Extraction and Compressed Sensing for Dimensionality Reduction

Anita Desiani[1,*], Sri Indra Maiyanti[2], Kanda Januar Miraswan[3], Muhammad Arhami[4]

[1,2,3]Departement of Mathematics, Mathematics and Science Faculty, Universitas Sriwijaya,
[4] Departement of Information and Computer Technology, Politeknik Negeri Lhokseumawe
**anita_desiani@unsri.ac.id*

ABSTRACT

Compressive sensing reduces the number of samples required to achieve acceptable reconstruction for medical diagnostics, therefore this research will implement dimensional reduction algorithms through compressed sensing for electrocardiogram signals (EKG). dimensional reduction is performed based on the fact that ECG signals can be reconstructed with linear combination coefficients with a bumpy base of small measurements with high accuracy. This study will use Principal Component Analysis (PCA) for feature extraction on ECG signals. The data used are the ECG patient records on the website page www.physionet.org as many as 1200 with each attribute as many as 256 attributes. The total data dimension used is 1200x256, which means the data has 1200 rows and has as many as 256 columns. To show the accuracy of the dimensional reduction result, so it is performed classification on data using KNN and Naive Bayes. The classification results show that KKN can classify well with 84.02% accuracy rate and the Naive Bayes accuracy is 65.78%. for 100 dimensions. The conclusion is those dimensional reductions for ECG data that have large dimensions, it still able to provide valid information like it uses the original data. Principle Component Analysis is a fit method for reducing data dimensions by selecting certain features, so the dimensions of the data become smaller but still able to provide well accuracy to the reader.

Keywords: Compressive Sensing, ECG signal, KNN, PCA

1. INTRODUCTION

Compressive sensing (CS) has become a rapidly expanding research area. Donoho and Candés et al., have proven if the signal own a representation of sparse in substructure, it will be reconstructed from few points of confinement [1], [2]. Some reconstruction approaches have been successfully performed therein. One approach that has been done is a dimensional reduction. The application of CS has been broadly implemented in various types of research such as Quasi-Incoherent Dictionaries, redundant dictionary. Signal Frequencies are related to the sparse signal representation[3]–[5]. Dimensional reduction is frequently recommended for generalizations and fewer features based on the number of original signals. The main purpose of dimensional reduction is to preserve the information from the original data provided as an optimization criterion or relevance so that there is a solution [6].

The Dimensional reduction applies to three stages of data , namely multivariate, analysis, and visualization[7]. some methods for data dimensional reduction are, multidimensional scaling (MDS)[8], Independent Independent Component (ICA)[9], and Principle Component and Analysis (PCA)[10]. Some of these methods are simple approach, but some methods also have limitations in strong linearity assumptions and new variables that are rather complex to address. The implementation of PCA involving decomposition of eigenvalues requires computation demand and if the dimension exceeds the amount of training data it results a singularity. Hence, to address these problems, we proposes the dimensional reduction in the three schemes, namely the first is use of PCA for feature extraction, the second is use compressed sensing algorithms for feature selection and the third is classifications using KNN classifier and Naive Bayes.

ECG (Electrocardiography) is the account of the electrical movement of the heart. It's in the type of transthoracic elucidation of electrical action of the heart over some stretch of time identified by anodes joined to the surface of the skin and recorded by a gadget outside to the body. The Chronicle created by this intrusive method is named an electrocardiogram. The outcomes from electrocardiogram recording are broadly used to distinguish different heart abnormalities[11]. ECG recording can't be utilized straightforwardly, instead, requires additional steps, for example, extraction from clamor amid recording, grouping, and examination, with the goal that the account results can be utilized to distinguish different heart ailments. A few analysts have attempted different techniques to break down the aftereffects of ECG recording, including utilizing the Wavelet change method[12]–[14].

Since the amount of these records can develop fundamentally so pressure is required for diminishing the capacity and transmission time. Compressive detecting has been effectively connected to ECG signals; decreasing the required number of tests to accomplish a worthy reproduction for therapeutic conclusion. In computerized examination catching notable highlights inside pre-determined districts of visual field issues happens i.e.; too expansive component brings about overfitting and poor speculation, expanding computational overhead while few element may not be adequate having absence of data. That is the reason dimensionality decrease approach is as reasonable. Compressive examination ECG bio-flag attributable to the considerable distinction between its rate of progress and its rate of data.

In this paper, we propose calculation compacted detecting dimensionality lessening for highlight choice. Traditional Nyquist inspecting in light of previous requires digitization at the rate no less than double the most extreme data transmission. Compacted detecting packs simple esteems that can be digitized and prepared at a decreased rate which spares the vitality utilizing CS. The better pressure is the mandatory advance which lessens the computational cost. CS has major favorable circumstances, for example, exchanging the computational overhead from encoder to decoder and area of the biggest coefficient on wavelet space should not be encoded.

2. IMPLEMENTATION OF DIMENSIONALITY REDUCTION ALGORITHM

2.1 COMPRENSIVE SENSING

Data can be represented as a sparse vector in the highest dimensional space for various applications, for example images in the wavelet domain [2][7], data bag model for text classification and natural language processing [5][6], network sensors

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

and communications [8], and data stream [4]. Compressive sensing is one of the Feature selection. CS can replace certain signals from smaller samples than needed in Nyquist samples. Recovery is appropriate if the signal has a low level of information (meaning the existing sparse comes from the original data or some domain transformation). The number of samples required for recovery depends on the particular reconstruction algorithm used previously. If the signal is not sparse, it will achieve the best reconstruction signal recovery obtained from the largest signal coefficient. CS handles noise and reconstruction errors that are restricted to data interruptions.

Compressive Sensing (CS) aims to recover signals with N-dimensional through the reception of a small sample of L, which is ideally much smaller than N. This study will show how data have sparse representation, and some unknown substructures use compressed sensing as an efficient one-to-one transformation[21]. The use of CS to provide $m \times n$ matrix A, with the smallest possible m measurement and recovery algorithm Δ so that for each signal k-sparse $x \in R_n$, Δ can recover x from the measurement $b = Ax$.

2.3 FEATURE EXTRACTION-PCA

PCA has features which can be applied from information extraction to dimensional reduction for data visualization[10]. The PCA used in this study will help to obtain the sparsity signal transformation required for CS to improve the signal through a good approach while the original signal of the sample with a small size. The approach used adapts to non-stationary signals in the real world through the online estimation of related properties in space and time; it is then utilized by the PCA to derive transformation for CS. We can have implementation of PCA steps as below:

PCA steps: Performed on D^l

Parameters: *K*:number of principle components or determined algorithmically
Initialization none particular
Inputx:list of data vectors (D^l), i=1,..n
1) Compute: μ:d-dim mean vector
2) Compute ∑:dxd covariance matrix
3)Selection of k largest Eigenvalues and corresponding Eigen vectors
4) Build dxk matrix A with column consisting of the k Eigen vectors
5)Projection of data x on to k-dim subspace x': x'=F(x)=A'(x-μ)
Output x': list of transformed data vector

Implementation of the PCA.DAT is size n x d. nObs=n

```
Coeff   = princomp(DAT);          %[nDim,nDim]
nPco= round(min(size(DAT))*0.7); % reduced dimension
PXO=coeff(:,1:nPco);   %select the 1stnPco eigenvector
DATRed=zeros(nObs,nPco);
For i=1:nObs
DATRed(I,: )=DAT(I,; )* PCO; %transform each sample end
```

In our experiment we reduce the dimensionality of given ECG signal which has 1200*256 and using PCA dimensional reduction algorithm we reduced the dimensionality to 10, 50, 80 and 100 (Figure 1, and Figure 2).

3. CLASSIFICATION

In our paper we consider to classification methods KNN classifier and Naive Bayes

3.1 K-NEAREST NEIGHBOR

The k-closest neighbor algorithm is among the most straightforward of all machine learning calculations. Given a testing set, we simply store all its samples as an exhaustive reference and to classify a testing sample, we compare all the training samples to it to arrive at a classification decision; that is we do not really ”relate” the training samples in any way. KNN classifiers a non-parametric method used for classification. The information comprises of the k nearest preparing cases in the element space. In KNN, objects are grouped by the proximity distance of their neighbors. the data object will be assigned in the most common class among its nearest neighbors. KNN is an algorithm based on an instance, or also called a lazy algorithm. The function at KNN is approached locally and all calculations are suspended until the classification process is completed so that the closer neighbors will contribute to each other in a class rather than the further neighbors. At the classification stage, a constant k is given, where k is a constant specified by the user and an instant vector that is an unknown label. the label on the vector is determined by assigning the most common label to the training sample or closest to the query point.

KNN classification. D^l=TRN D^T=TST. G with group label	
Initialization	normalize data
Training	training samples D^L with class label G
Testing	for testing sample : compute distances to all training sample→ D rank order D →D^r
Decision	observe the 1st ranked distances in D^r (the k nearest neighbor

In our experiment we consider the k values as 1, 3,5,7,9 and do the classification having the train and test data in hand. Which is 1200x256 and we even experiment the classification after reducing the dimensionality to 10,50,80, 100. Figure 3, 4 and 5 are showing the example of classification with differeng of K in KNN.

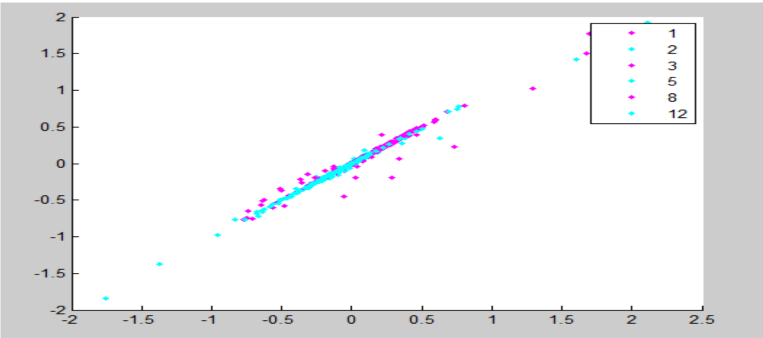

FIGURE 1. Scatter diagram show classification for Train data and test data with 1200x256 for K=1

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

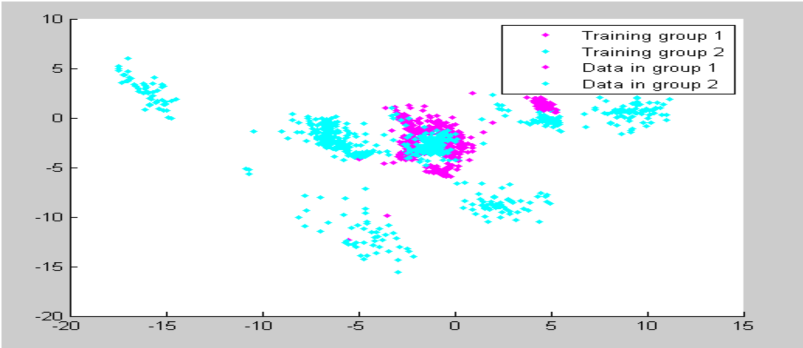

FIGURE 2. Scatter diagram show classification *for dimensionality 10 and k=3*

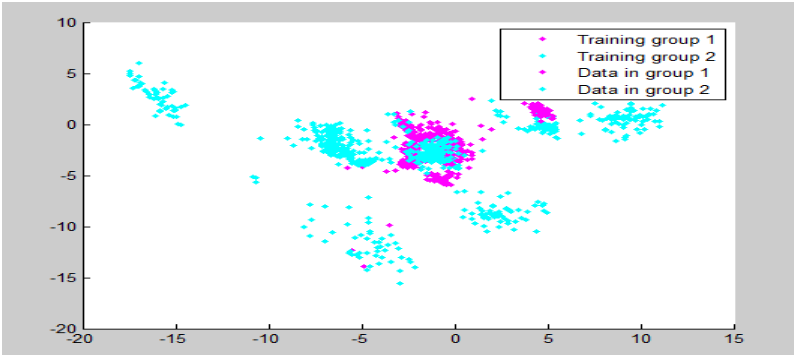

FIGURE 3. Scatter diagram show classification *for dimensionality 50 and k=3*

3.2 NAÏVE BAYES

This paper shows a strategy for ordering patients as indicated by measurable highlights separated from their ECG signals utilizing a hereditarily advanced Bayesian classifier. It accepts a concept probabilistic model and it enables to catch vulnerability about the mainly model by determining probabilities of the outcomes. It can take care of analytic and prescient issues. It is utilized the information on earlier occasions to anticipate future occasions. Naive Bayesian characterization depends on the Bayesian probabilistic it is notably, suitable when the dimension of the information sources is high. Parameter estimation for the Naive Bayes exploits the strategy for most extreme possibility. It is better for the complex real problem.

The Naive Bayes classifier, one goes a step further and even makes the decision based on density estimation and this classifier is thus theoretically the most elegant model, as everything is based on parameterization. Practically, the classifier has only limited success, as too much elegance sometimes lacks the robustness to deal with 'messy' data. The Naive Bayes classifier performs density estimation assuming uni-modal Gaussian distributions, namely by taking the mean and the standard deviations of the individual feature dimensions for each class (group).

Naïve Bayes Classifier k=1,.....c
Training c classes ($\in D^L$) Mean μ_k , covariance $\sum_k$ determine $\lvert\sum_k\rvert$ inverse $\sum_k - 1$ Prior P(k)→gk Testing 1)for a testing sample x$\in D^T$ determine g(x) for all c classes→ g_k 2)multiply each g_k with the class prior P(k): $f_k = g_k$ ..P(k) Decision choose maximum of f_k : argmaxk f_k

Favorable position of Naïve Bayes is it just requires a little measure of preparing information to appraise the parameters (means and differences in the factors) important for order. Since autonomous factors are accepted, just the differences of the factors for each class should be resolved and not the whole covariance framework. To know which classifier performance is better we compute its accuracy as given below in Table we can clearly see that KNN is more accurate.

TABLE 1.
The Comparison of Naive Bayes and KNN

Method	Accuracy (%) Dimensional				
	256	10	50	80	100
Naïve Bayes	67.67	55.58	56,08	65.17	66,78
KNN	86.33	60.75	68.92	85.91	87,02

TABLE 2.
KNN based Accuracy Having Different Values of K

Method	Accuracy (%) Dimensional				
	256	10	50	80	100
K=1	86.33	60,75	68.92	85.91	87,02
K=3	86.33	60.92	68.25	85.90	87,02
K=5	86.33	60.17	68.25	85.93	87, 11
K=7	86	60.33	68.24	85.75	87
k=9	85.83	60.75	68.42	85.8	87

From Table 1 and 2, we can see the KNN method is better than Naive Bayes. The best accuracy for KNN method can achieve on 100 dimensional and for k=5. It shows that PCA is able to reduce number of dimensional data but it still has a good accuracy of KNN method even it used less dimensional data. The PCA can be reduce memory.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

4. CONCLUSION

In this paper, we discussed another approach for the storage of ECG signals based on compressed sensing for general and walking or symptomatic quality observing and we also discussed dimensional reduction and classification. Here we have presented method for feature extraction PCA- based dimensionality reduction which decreases the dimensionality of information by choosing just a subset of calculate feature to build a model. The classification algorithm selects discriminatory features and classifies those features as training data simultaneously. By reducing dimensionality we not only reduced the complexity of computation but time and save memory but it still has a good accurary for classification.

REFERENCES

[1] D. L. Donoho, "Compressed Sensing," *IEEE Trans. Infor. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[2] L. F. G. T, E. D. T, J. C. Ria, and D. Castellanos, "Feature Selection using Hybrid Evaluation Approaches based on Genetic Algorithms," *IEEE Trans. Infor. Theory*, vol. 1, pp. 2–7, 2006.

[3] R. Gribonval and P. Vandergheynst, "On the Exponential Convergence of Matching Pursuits in Quasi-Incoherent Dictionaries," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 360–365, 2006.

[4] E. J. Candès, J. Romberg, and T. Tao, "Robust Uncertainty Principles : Exact Signal Frequency Information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[5] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed Sensing and Redundant Dictionaries," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 2210–2219, 2008.

[6] E. Delgado-trejos and A. Perera-lluna, "Dimensionality Reduction Oriented Toward the Feature Visualization for Ischemia Detection," in *IEEE Comput. Soc. Press Proc., Electron., Robot. Automotive Mech. Conf (CERMA)*, 2009, vol. 13, no. 4, pp. 590–598.

[7] A. Konig, "Dimensionality Reduction Techniques for Multivariate Data Classification , Interactive Visualization , and Analysis - Systematic Feature Selection vs . Extraction," in *4th Int. Conf. Knowl.-Based Intell. Eng. Syst. Allied Technol., Brighton, U.K.*, 2000.

[8] T. F. Cox and M. A. A. Cox, *Multidimensional Scaling*, 2nd ed. London: Chapman & Hall/CRC Press, 2000.

[9] E. O. A. Hyvarinen, J. Karhunen, *Independent Component Analysis (Wiley Series on Adaptive and Learning Systems for Signal Processing, Communications and Control)*. New York: Wiley, 2001.

[10] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. New York: Spring-Verlak, 2002.

[11] Acharya U. Rajendra, J. Suri, J. Spaan, *Advances in Cardiac Signal Processing*. New York, NY, USA: Springer-Verlag, 2007.

[12] A. Navaria, "Denoising and Feature Extraction of ECG," *Int. J. Emerg. Technol. Comput. Appl. Sci. (IJETCAS)*, vol. 13, pp. 222–226, 2013.

[13] N. Emanet, "ECG Beat Classification by Using Discrete Wavelet Transform

and Random Forest Algorithm," in *Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control, ICSCCW, Fifth International Conference on Cite this publication Cite this publication*, 2009.

[14] A. K. Manocha and M. Singh, "Automatic Delineation of ECG Characteristics Points using Window Search & Multi-resolution Wavelet Transform approach," in *Proc. of Int. Conf. on Emerging Trends in Engineering and Technology, ACEE*, 2013.

[15] van de M. L.J.P, "An Introduction to Dimensionality Reduction Review," Netherlands, 2007.

[16] C.J.C. Burges, "A Complete Guide for Practitioners and Researchers, chapter Geometric Methods for Feature Selection and Dimensional Reduction," in *Data Mining and Knowledge Discovery Handbook*, Netherlands: Kluwer Academic Publishers, 2005.

[17] Scholkopf, Bernhard., Alexander Smola and KlausRobert Muller, "Nonlinear Component Analysis as a Kernel Eigenvalue Problem," *Neural Comput.*, vol. 1319, pp. 1299–1319, 1998.

[18] J. Wang and Z. Zhang, "Adaptive Manifold Learning," The MIT press, Cambridge, 2005.

[19] I. Daubechies, *No TitleTen Lectures on Wavelets*, vol. 16. 1992.

[20] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, 2003.

[21] M. Stephane, *A Wavelet Tour of Signal Processing*. London: AP Professional, 2008.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)